

FEATURE EXTRACTION FROM HYPERSPECTRAL IMAGES COMPRESSED USING THE JPEG-2000 STANDARD

Mihaela D. Quirk

Los Alamos National Laboratory and
The University of Texas at Austin
pal@lanl.gov

Christopher M. Brislawn

Los Alamos National Laboratory
Los Alamos NM 87545
brislawn@lanl.gov

Steven P. Brumby

Los Alamos National Laboratory
Los Alamos NM 87545
brumby@lanl.gov

Abstract

We present results quantifying the exploitability of compressed remote sensing imagery. The performance of various feature extraction and classification tasks is measured on hyperspectral images coded using the JPEG-2000 Standard. Spectral decorrelation is performed using the Karhunen-Loève Transform and the 9-7 wavelet transform as part of the JPEG-2000 process. The quantitative performance of supervised, unsupervised, and hybrid classification tasks is reported as a function of the compressed bit rate for each spectral decorrelation scheme. The tasks examined are shown to perform with 99% accuracy at rates as low as 0.125 bits/pixel/band. This suggests that one need not limit remote sensing systems to lossless compression only, since many common classification tools perform reliably on images compressed to very low bit rates.

1. Introduction

The multicomponent coding features of the JPEG-2000 Standard [1, 2] make efficient lossy compression of hyperspectral imagery a viable alternative to lossless compression methods. The only studies to date on the exploitability of lossily compressed and reconstructed hyperspectral imagery, however, are quite recent and few in number. The papers by Brower *et al.* [3] and Kasner *et al.* [6] compared different combinations of JPEG-2000 options on hyperspectral imagery in terms of root mean-square error or peak signal to noise ratio. In Shen and Kasner [10], the authors studied anomaly detection and material identification tasks. They reported exploitation task performance on JPEG-2000-compressed imagery in terms of the number of targets correctly detected or identified at a fixed false alarm rate. We expand on Shen and Kasner's approach by considering a number of representative remote-sensing feature extraction tasks performed on imagery compressed to a wide range of bit rates.

1.1. The JPEG-2000 Standard

The new still image compression standard, JPEG-2000, was written with explicit requirements for supporting coding and compression of multicomponent imagery. The standard is equipped with a variety of options for transforming and manipulating collections of image components (e.g., spectral bands). A high level overview of the JPEG-2000 Standard is presented in Figure 1. Component decorrelation transforms are followed by spatial wavelet transforms, rate allocation and quantization, and binary arithmetic bit-plane encoding. The compressed bitstreams are signaled in packets that can be ordered according to a variety of possible priorities to support various progressive transmission objectives. For instance, the initial portion of a truncated JPEG-2000 codestream can always be decoded to yield an approximation of the image represented by the full codestream. This feature, known as an *embedded codestream*, enables multiple users to access the same compressed codestream at a variety of different levels of fidelity, which is useful in database applications. A detailed technical presentation of the JPEG-2000 standard can be found in Taubman and Marcellin [11].

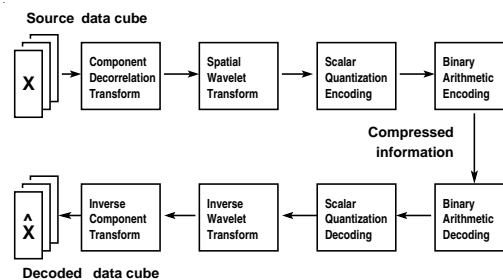


Figure 1. High level overview of JPEG-2000.

In the experiments below, an AVIRIS hyperspectral image is compressed and reconstructed using JPEG-2000 at bit rates varying from 0.125 bits/pixel/band (bppb) to 4 bppb. AVIRIS data is highly correlated along the spectral axis, and we exploit this fact with 2 different component decorrelation transforms. The fidelity of reconstructed images is quantified and reported as a function of bit rate.

For instance, one simple way of quantifying fidelity is to

report the SNR of a reconstructed image cube (“3-D SNR”), where “truth” is given by the original, uncompressed image cube. Typical rate-distortion performance for such measurements is shown in Figure 2. The curves present 3-D SNR as a function of rate for 3 component transform options: no component decorrelation, wavelet transform decorrelation, and KLT decorrelation. The other steps in the JPEG-2000 process were identical in all 3 cases. Observe how (nonadaptive) 9-7 wavelet transform decorrelation yields a gain of around 10-12 dB on this particular image, while image-dependent KLT decorrelation produces around 15-20 dB of gain. Results from the following experiments are reported in a similar fashion, with 3-D SNR replaced by metrics based on exploitation task performance.

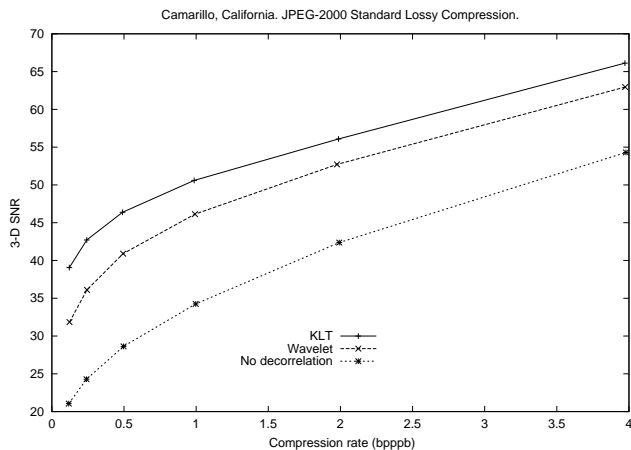


Figure 2. 3-D SNR for the “Camarillo” image.

1.2. Data Used in Experiments

The experiments were performed on Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) hyperspectral images. AVIRIS is a JPL instrument that delivers calibrated images of the upwelling spectral radiance in 224 contiguous spectral bands with wavelengths from 400 to 2500 nm with a spatial resolution varying from a few meters to 20 meters. Each spectral component has 512 x 614 pixels with a sample precision of 16 bpppb [12]. The “Clouds” scene (Figure 3) is used in atmospheric correction problems. “Moffet Field, California” (Figure 6), has a mix of roads, urban areas, and vegetation, and is suited for spatial-spectral image content analysis. The “Camarillo, California” scene (Figure 10) has been used for monitoring vegetation and for developing differential absorption techniques to retrieve columnar water vapor [8].

Training Data. Many classification tasks employ training data, often designated as “true” and “false” classes. “True” classes may be derived from actual ground truth, collected on the ground at the time the image was acquired, or they may be determined by an analyst from the images. In our experiments the “true” and “false” classes were determined by analysts with the aid of a software tool called

ALADDIN [4]. An example training class is shown for road features in the Moffet Field image in Figure 7.

2. Feature Extraction Tasks and Results

We examine the performance of 5 classification tasks and a hyperspectral feature transform on reconstructed imagery. There are three main types of classification for remote sensing data [9, 7]: supervised, unsupervised, and hybrid. In supervised classification, one or more training classes are supplied to determine the classes constituting a thematic map. The classifier maps every pixel into one of the desired feature classes by means of spectral analysis. The supervised classifications studied here are: Spectral Angle Mapping (SAM), Binary Encoding, and Minimum Distance classification.

The only unsupervised classification task examined is K-means clustering. The process is unsupervised in the sense that no prior information on feature classes is given. The image is segmented automatically into Voronoi cells, and the user only supplies the desired number of classes and convergence conditions. Classification performance of supervised and unsupervised tasks is reported as percentage of correctly classified pixels, i.e., percentage of pixels in the reconstructed image whose classification is the same as in the original image.

A hybrid supervised-unsupervised classifier starts with an analyst’s input but performs tasks automatically. GENIE [4] is a hybrid algorithm that processes data both spatially and spectrally. GENIE performance is quantified in terms of a fitness metric presented below. In addition, we compute a derived hyperspectral image feature known as the “Normalized Difference Vegetation Index” (NDVI), which is a scalar field defined at each pixel in the image. The fidelity of NDVI computations is defined as the SNR of the 2-D NDVI field. Except for GENIE, all experiments were performed using procedures supplied with the ENVI software package [5].

2.1. Spectral Angle Mapper Classification

The Spectral Angle Mapper (SAM) computes the normalized inner product of training pixels with image pixels and assigns pixels to a feature class if the angle is less than a user-supplied threshold. SAM is insensitive to pixel magnitudes and is suited for analyzing scenes with changes in illumination, shadows, etc. Figures 4 and 5 show a thematic map and quantitative SAM performance for the cloud classification problem.

2.2. Minimum Distance Classifier

The minimum distance algorithm classifies pixels using a training class (in this case, the “road” class in Figure 7). A pixel is considered in-class if the Euclidean distance from the pixel vector to the class mean vector is less than a user-supplied threshold. Performance on the “roads” feature class in Moffet Field is shown in Figure 8.

2.3. Binary Encoding

Binary Encoding quantizes a pixel vector to a binary vector by comparing each sample to that vector's mean sample value. A sample is quantized to 1 if it lies above the vector mean, 0 if it lies below the mean. This procedure is applied to each data vector being classified and to the mean vector from a training class. The Hamming distance between the quantized data vectors and the quantized reference vector can then be calculated very efficiently using a Boolean exclusive OR operation. Classification of a pixel to a training class is based on a user-supplied threshold for the Hamming distance. Figure 9 displays results from Binary Encoding.

2.4. K-Means Clustering

K-Means unsupervised classification starts with evenly distributed initial reference vectors, one per desired feature class. Data pixels are clustered using a minimum distance criterion. The process then iteratively recalculates class means and reclassifies pixels with respect to the new means. The process stops when the number of pixels in each class changes by less than a specified threshold or when a limit on the number of iterations is reached. A thematic map for 5-class K-means classification and quantitative results are shown for Camarillo data in Figures 11 and 12.

2.5. GENIE Hybrid Evolutionary Classification

GENIE (GENetic Imagery Exploitation) is a hybrid method. The evolutionary part of the program selects a set of image processing operations that transform raw image planes into new components. These intermediate feature components are then input to a conventional supervised classification technique. The *fitness*, F , for a candidate solution is defined as $F = 500(R_d + (1 - R_f))$. R_d is the detection rate, and R_f is the false alarm rate. A fitness of 1000 indicates a "perfect" classification result, i.e., none of the pixels have been classified incorrectly. Fitness results for "road" classification of Moffett Field data are shown in Figure 13.

2.6. Normalized Difference Vegetation Index

The NDVI is a ratio between the difference and the sum of near-infrared and red bands. It is employed as a feature in classification tasks used to monitor vegetation. In this experiment, rather than quantify the performance of an NDVI-based classification task we report the 2-D SNR of the NDVI field derived from reconstructed Camarillo data. As shown in Figure 14, NDVI SNR gains anywhere from 5-12 dB when using wavelet component decorrelation (versus no component decorrelation), while KLT decorrelation yields an additional 2-4 dB.

3. Discussion/Conclusion

All the feature extraction tools investigated render over 99.99% correctly classified points at 4 bpppb, a bit rate

that cannot be obtained using lossless compression. Decorrelation in the component direction prior to compression gives a superior percentage of correctly classified points when compared to classification with no component decorrelation. This behavior is consistent for all the classification tasks considered. For large scale applications, coding hyperspectral images with JPEG-2000 at 0.125 bpppb using image-dependent KLT component decorrelation allows users to transmit less than 1% of the original data while still maintaining 99.5% classification accuracy. With nonadaptive wavelet transform component decorrelation, classification accuracy at 0.125 bpppb is only slightly worse, around 98%.

References

- [1] *JPEG 2000 Image Coding System, Part 1*, ISO/IEC Int'l. Standard 15444-1, ITU-T Rec. T.800. Int'l. Org. for Standardization, Dec. 2000.
- [2] *JPEG 2000 Image Coding System, Part 2 (Extensions)*, ISO/IEC Int'l. Standard 15444-2, ITU-T Rec. T.800. Int'l. Org. for Standardization, Dec. 2001.
- [3] B. V. Brower, A. Lan, J. Kasner, and S. Shen. Multiple component compression within jpeg-2000 as compared with other techniques. *Applications of Digital Image Processing XXIII, Proceedings of SPIE*, 4115:544-551, 2000.
- [4] N. R. Harvey, S. P. Brumby, S. Perkins, J. Theiler, J. J. Szymanski, J. J. Bloch, R. B. Porter, M. Gallassi, and A. C. Young. Image feature extraction: Genie versus conventional supervised classification techniques. *IEEE Transactions on Geoscience and Remote Sensing*, 40, 2002. To appear.
- [5] <http://www.rsinc.com/Env/tut2.cfm>.
- [6] J. Kasner, A. B. M. W. Marcellin, A. Lan, B. V. Brower, S. S. Shen, and T. S. Wilkinson. JPEG-2000 compression using 3D wavelets and KLT with application to HYDICE data. *Imaging Spectrometry VI, Proceedings of SPIE*, 4132:157-166, 2000.
- [7] J. A. Richards and X. Jia. *Remote Sensing Digital Image Analysis - An Introduction*. Springer-Verlag, Berlin Heidelberg, 1999.
- [8] D. Schlappfer, C. Borel, J. Keller, and K. I. Itten. Atmospheric pre-corrected differential absorption techniques to retrieve columnar water vapor. *Remote Sensing of Environment*, 65(3):353-366, 1998.
- [9] R. A. Schowengerdt. *Remote Sensing Models and Methods for Image Processing*. Academic Press, second edition, 1997.
- [10] S. S. Shen and K. H. Kasner. Effects of 3D wavelets and KLT based JPEG-2000 - hyperspectral compression on exploitation. *Imaging Spectrometry VI, Proceedings of SPIE*, 4132:167-176, 2000.
- [11] D. S. Taubman and M. W. Marcellin. *JPEG-2000 : Image Compression Fundamentals, Standards, and Practice*. Kluwer International Series in Engineering and Computer Science, 2001.
- [12] G. Vane, A. F. H. Goetz, and J. B. Wellman. Airborne imaging spectrometer: A new tool for remote sensing. *Proc. 1984 IEEE Transactions on Geoscience and Remote Sensing*, GE-22(6):546-549, 1984.

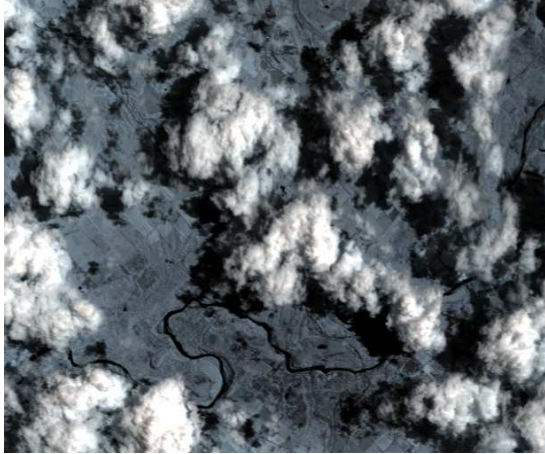


Figure 3. The “Clouds” scene.

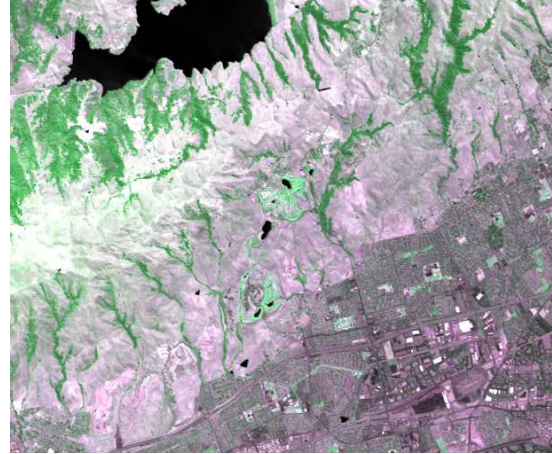


Figure 6. The “Moffet Field” image.

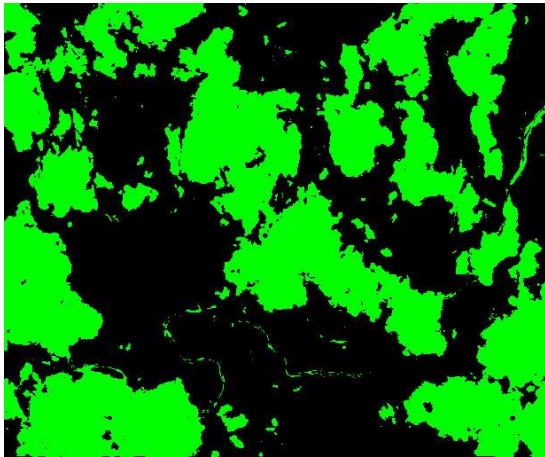


Figure 4. SAM thematic map for the “cloud” class.



Figure 7. Moffet Field training class for “road” features.

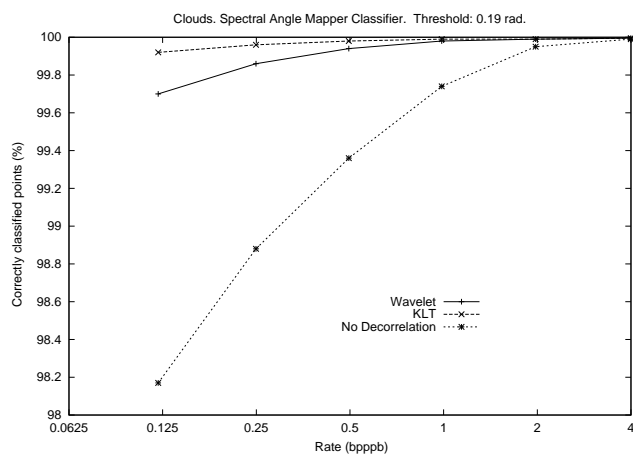


Figure 5. Percentage of correctly classified “cloud” pixels with Spectral Angle Mapper.

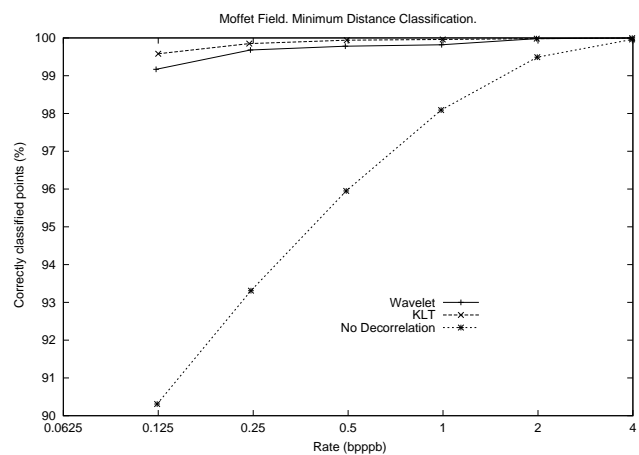


Figure 8. Performance of Minimum Distance classification for the “road” class.

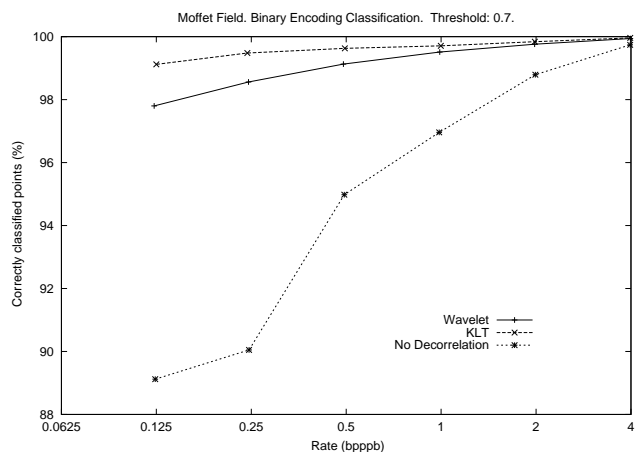


Figure 9. Performance of Binary Encoding classification for the “road” class.

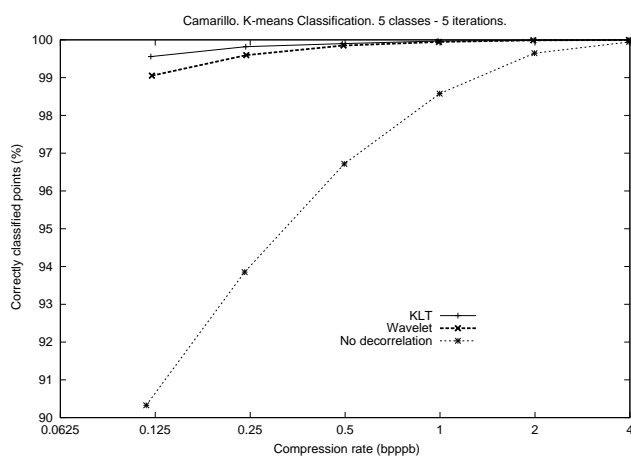


Figure 12. K-means classification performance on “Camarillo.”

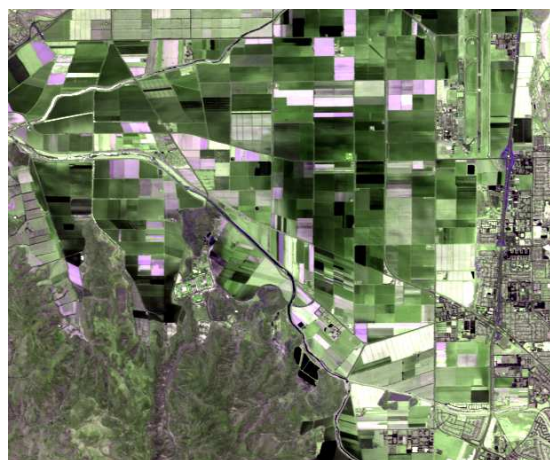


Figure 10. The “Camarillo” image.

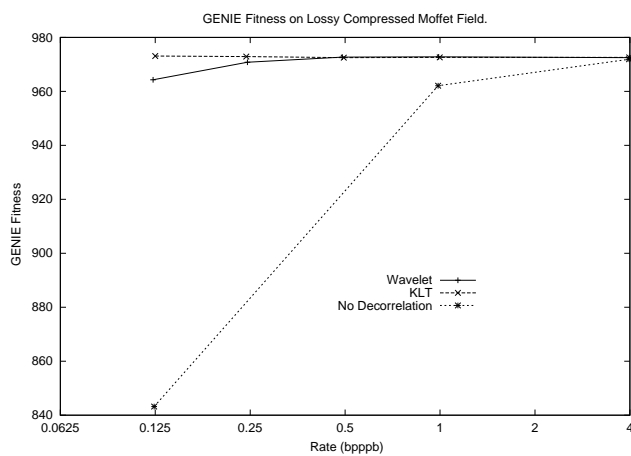


Figure 13. GENIE fitness for “road” class.

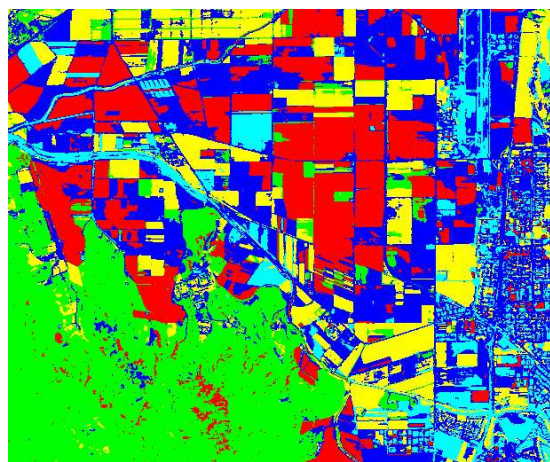


Figure 11. K-means 5-class thematic map for “Camarillo.”

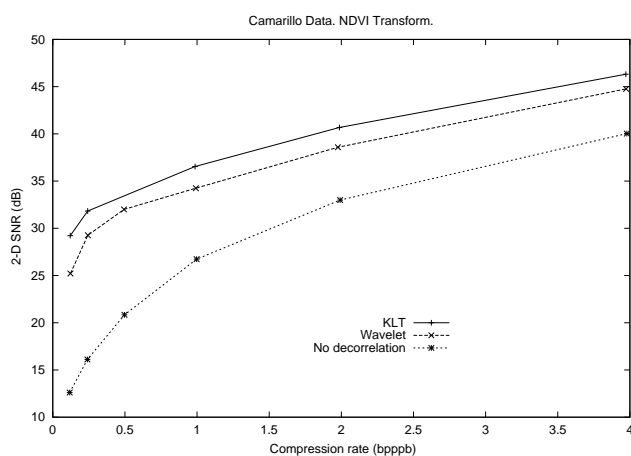


Figure 14. SNR for NDVI feature computation on “Camarillo.”